PH661 – Fall 2008

Statistical methods for physics and astrophysics

Assignment #3 – Thursday, Sept 4, 2008

1. Consider the random variable

$$\bar{x} = \frac{1}{N} \sum_{i=1}^{N} x_i \tag{1}$$

where x_i are identical independent random variables with mean μ and variance σ^2. Show that the variance of $\bar{x}$ is equal to σ^2/N.

Solution:

Need to make use of the assumption that the variables are identically distributed, and independent. In that case, first use a fundamental property of the variance, $Var[aX] = a^2 Var[X]$:

$Var[\bar{x}] = Var[\frac{1}{N}\sum_{i=1}^{N} x_i] = \frac{1}{N^2} Var[\sum_{i=1}^{N} x_i]$

Then make use of the fact that variances add for independent variables, $Var[\sum_{i=1}^{N} x_i] = N Var[x] = N\sigma^2$,

$\Rightarrow Var[\bar{x}] = \sigma^2/N$.

2. It is known that the I.Q. scores are distributed like a Gaussian, with $\mu = 100$ and $\sigma^2 = 16$. Assume that the sum of Gaussian variables is also distributed like a Gaussian.

(a) What is the probability that the *mean I.Q.* of a random group of 21 people will be between 95 and 105:

Solution:

Start by considering that the mean of N measurements is distributed like a Gaussian with the same mean as the distribution of each measurement, but with a smaller variance, equal to $\sigma_\mu^2 = \sigma^2/N = 16/21 = 0.76$

The interval $[95, 105]$ therefore corresponds to $[\mu - 6.6\sigma, \mu + 6.6\sigma]$ which corresponds to a probability in excess of 99.99999% (it is off the chart in Table C.2).

(b) How large a sample is required in order to have a chance of 95% that the *mean I.Q.* of the group lies between 95 and 105?

Solution:

In this case, a 95% probability is included in an interval $[\mu - 1.96\sigma, \mu + 1.96\sigma]$, thus the standard error of the mean must be such that:

$1.96 \cdot \left(\sigma/\sqrt{N}\right) = 5 \Rightarrow N = \left(\frac{1.96 \times 4}{5}\right)^2 = 2.5,$

so one could choose a sample of 2 or 3 people to approximate the requirement.

Notice that each score has a probability $\simeq 77\%$ of being in the $[95, 105]$ interval.

3. The frequency of twins in European population is about 12 in every 1,000 maternities. What is the probability that there are no twins in 200 births, using (a) the binomial distribution, and (b) the Poisson distribution?

Solution:

(a) Binomial with p=0.012, q=0.988, N=200.

A={no twins in 200 births}

P(A)=$P_N(0) = 0.012^0 \cdot 0.988^{200} = 0.0894$

(b) Use Poisson (μ) with

$\mu = N \cdot p = 2.4$

P(A)=$P(0) = \frac{2.4^0}{0!} \cdot e^{-2.4} = 0.091$

Poisson approximation seems a good approximation (to within 2%).